

Análise dos Classificadores Bayesianos utilizando dados hiperespectrais Hyperion / EO1

Alan Bernardo Arruda Bisso^{*}
Laurindo Antonio Guasselli^{**}
Dejanira Luderitz Saldanha^{***}

Resumo

A região nordeste do estado do Rio Grande do Sul abriga extensos conjuntos florestais do Bioma Mata Atlântica. Essas áreas se encontram extremamente fragmentadas e com reduzida parcela de vegetação nativa. O cultivo de banana ocupa extensas áreas principalmente nas encostas, anteriormente vegetadas, gerando um enorme impacto sobre as áreas de Mata Atlântica. O presente trabalho avalia o uso de classificadores bayesianos com dados hiperespectrais para a diferenciação das áreas de Mata Atlântica e de bananais no município de Três Cachoeiras, o maior produtor de banana do Estado. O melhor resultado encontrado utiliza o método Máxima Verossimilhança com limiar de exclusão de 2% , exatidão global de 0,77 e índice Kappa de 0,48. No método Distância Euclidiana foi encontrado o pior resultado com exatidão global de 0,59 e índice Kappa de 0,22.

Palavras-chave: Hiperespectral; Mata Atlântica; Classificação.

Analysis of Bayesian classifiers using hyperspectral Hyperion / EO1 data

Abstract

The Northeastern region of Rio Grande do Sul is home to sets of Atlantic Forest in the state; however they are extremely

* Mestre em Sensoriamento Remoto pela UFRGS.

** Doutor em Recursos Hídricos e Saneamento Ambiental pela UFRGS.

*** Pós-Doutorado na Universidade de Coimbra.

fragmented, with small portion of native vegetation. Banana plantation in this region is takes the biggest part of this area, and the features of this planting generates a huge impact on the areas of Atlantic Forest, since that it is giving way to the areas of plantations. This assignment aims to evaluate the use of Bayesian classifiers with hyperspectral data for differentiation between the Atlantic Forest and banana's areas in the town of Três Cachoeiras, the state's largest producer of banana. The best result was found with the Maximum Likelihood method with exclusion threshold of 2%, the overall accuracy was 0.77 and index Kappa was 0.48. In the Euclidean Distance method was found the worst result with an overall accuracy of 0.59 and Kappa 0.22.

Key words: Hyperspectral; Atlantic Forest; Classification.

Introdução

A região nordeste do Estado do Rio Grande do Sul apresenta as maiores áreas de remanescentes florestais da Mata Atlântica no estado. Nessa região a Floresta Ombrófila Densa ocorre predominantemente em áreas de relevo acidentado da faixa de transição dos compartimentos geomorfológicos da Planície Costeira e do Planalto Meridional. Os remanescentes da Mata Atlântica ocupam principalmente as áreas de morros e encostas íngremes e relevo mais acidentado, onde está parcialmente preservada. Nessas encostas a atividade agrícola é dificultada pela acentuada declividade, impedindo a mecanização, a fixação do solo e o desenvolvimento de outras culturas agrícolas.

A produção de banana constitui-se em uma das poucas culturas agrícolas que se adaptado bem nesta situação de relevo. Por serem perenes, os bananais vem sendo cultivados nos locais originalmente ocupados pela Mata Atlântica. Esta atividade permite a sobrevivência econômica de várias famílias da região, que é a maior produtora de banana no estado.

O sensoriamento remoto, a partir de imagens de satélites, permite o acompanhamento de mudanças na superfície da terra de maneira abrangente e em diversas escalas de análise, possibilitando monitorar mudanças ao longo do tempo e ao longo do espectro eletromagnético.

Os sensores multiespectrais, mais tradicionais, permitem uma diferenciação parcial das formações vegetais arbóreas com base em critérios espectrais, já que as pequenas diferenças ao longo do espectro são imperceptíveis em função das poucas bandas espectrais. Estas bandas abrangem intervalos largos ao longo do espectro eletromagnético, não sendo possível identificar feições espectrais mais estreitas, que diferenciam os alvos semelhantes. Além disso, as diferenciações texturais dos alvos são dificultadas, já que as resoluções espaciais são pouco refinadas, tornando difícil a caracterização das pequenas variações existentes, como: o espaçamento entre as plantas dos bananais e os diferentes arranjos e diversidade da composição das áreas de Mata Atlântica.

No sensoriamento hiperespectral, o processo de aquisição de imagens em centenas de bandas registradas e contíguas, possibilita que para cada pixel das mesmas seja possível derivar uma curva de reflectância espectral completa, denominadas de ‘espectroscopia de imageamento’ (*imaging spectroscopy*) ou ‘sensoriamento remoto hiperespectral’ (*hyperspectral remote sensing*). O objetivo do sensoriamento remoto hiperespectral é medir, quantitativamente, a assinatura espectral dos componentes do sistema Terra, a partir de espectros calibrados, adquiridos na forma de imagens, para uso em aplicações de sensoriamento remoto (INFOGEO, 2004). As imagens hiperespectrais permitem a confecção de espectros de reflectâncias de uma forma “quase” contínua, o que permite a caracterização dos elementos que formam a paisagem.

O sensor Hyperion/EO-1 é o primeiro sensor hiperespectral em nível orbital, adquirindo dados de 400 a 2500 nm em bandas estreitas e contíguas de 10 nm de resolução espectral e 30 m de resolução espacial. Têm sido utilizado em análises como: detecção de formações florestais (VIBRANS & SCHRAMM 2007), relação

entre índices de vegetação e produtividade de soja (GALVÃO, FORMAGGIO, BREUNIG, 2009) identificação de alvos agrícolas (TISOT *et al.*, 2007).

As diferentes classes de cobertura terrestre têm propriedades de reflectância distintas, porém algumas destas classes possuem propriedades de reflectância que se assemelham ou se sobrepõem, dificultando a separação das mesmas. Dessa forma, a maioria dos métodos de classificação se propõe a avaliar a semelhança de padrões espectrais usando alguma medida da distância entre pixels do espaço espectral (NOGUCHI, 2004), visam reconhecer padrões e alvos homogêneos com o objetivo de mapear os diferentes componentes da superfície terrestre (PORTILLO, 2007).

Os classificadores de Bayes baseiam-se nos histogramas das classes para aproximar funções de densidade de probabilidade que seguem o padrão normal de distribuição, conhecido como distribuição gaussiana. O presente trabalho tem como objetivo avaliar o uso de dados hiperespectrais do sensor Hyperion aplicando classificadores bayesianos: Máxima verossimilhança, Distância Euclidiana, Linear de Fisher com e sem a utilização de limiar de exclusão (*Threshold*).

Área de estudo

A área de estudo está localizada na região do Litoral Norte do Rio Grande do Sul, mais precisamente o Noroeste do município de Três Cachoeiras. Nessa região ocorre grande concentração de plantações de banana, predominante nas encostas dos morros (Figura 1). Essas áreas se caracterizam por encostas íngremes, cobertas anteriormente por Mata Atlântica, com orientação do relevo bastante diversa, o que dificulta a classificação devido ao intenso sombreamento das encostas. A área foi selecionada com base na disponibilidade de imagens Hyperion / EO-1 em que houvesse o imageamento de áreas de bananais e de Mata Atlântica. O recorte espacial da imagem esta definido pelas coordenadas 50° 01'O e 29° 29'S e 50° 05'O e 29° 24'S.

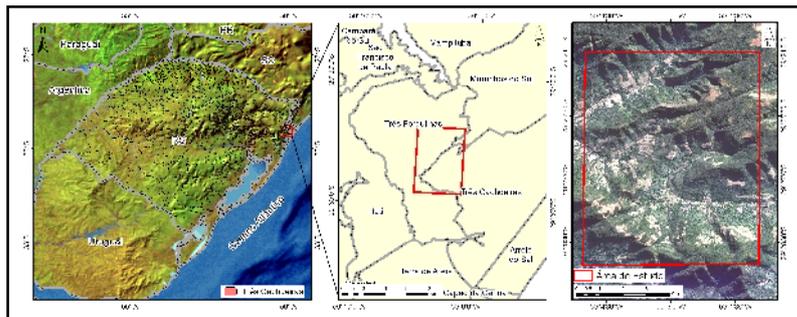


FIGURA 1: Localização da área de estudo, noroeste do município de Três Cachoeiras-RS, imagem Hyperion / EO-1

Metodologia

Para atingir os objetivos propostos foi utilizada uma imagem Hyperion - EO1, de 02 de maio de 2004, obtida gratuitamente no site da USGS, através do GLOVIS (*Global Visualization Viewer*), (<http://glovis.usgs.gov>), e uma imagem Ikonos de 07 de abril de 2004.

A imagem Hyperion possui 220 bandas espectrais de 0,4 a 2,5 micrometros e resolução espacial de 30 metros, pode imagear uma área de 7,5 x 100 km, fornecendo um mapeamento espectral detalhado em todas as 220 bandas com alta precisão radiométrica.

Sobre a imagem Hyperion, a partir do software ENVI 4.5, foram executados os seguintes pré-processamentos: (a) conversão dos dados de L1R para o formato padrão do ENVI; (b) georeferenciamento da imagem utilizando um arquivo do tipo “.met”; (c) correção dos *stripes*; e (d) geração de um arquivo com os fatores de escala para a correção atmosférica. Para isso foi utilizado um *plugin* disponibilizado no site da *ITT Visual Information Solutions*, (www.ittvis.com), denominado de *Hyperion Tools*.

A correção atmosférica foi realizada no módulo *FLAASH* (*Fast Line-of-sight Atmospheric Analysis of Spectral Hypercubes*)

do ENVI 4.5. Este módulo é baseado nos modelos de transferência radiativa MODTRAN-4 (*Moderate Resolution Transmittance Code*), (FELDE et al., 2003). Essa correção elimina os efeitos de espalhamento e absorção da radiação, ocasionados pela interação com a atmosfera.

O recorte da imagem priorizou as áreas com concentração de plantio de bananas e de ocorrência de Mata Atlântica. É uma região de vales a oeste do município de Três Cachoeiras. Para caracterização espectral foram definidas quatro classes de uso e ocupação do solo: Mata Atlântica, Banana, Campo e Sombra. As amostras (pixels) selecionadas de cada classe foram utilizadas no treinamento dos classificadores para definição dos parâmetros de similaridade espectral.

Foi realizado um trabalho de campo entre os dias 27/08/2009 e 30/08/2009, para verificar as características da área e identificar os alvos a serem estudados. A diferença entre a data da imagem e o trabalho de campo se justifica, já que, a cultura da banana é perene e não ocorreram grandes alterações nas áreas plantadas no período.

O total de pixels utilizados nos processos de classificação e seleção de feições foi de 511 pixels, distribuídos nas quatro classes de uso e ocupação do solo (Tabela 1).

TABELA 1: Classes de uso do solo e número de pixels das amostras

Classes de Uso e cobertura do Solo	Nº de pixels
Mata	129
Banana	135
Campo	51
Sombra	196

A seleção das amostras dos pixels foi realizada com base no reconhecimento de campo e na verificação da mistura dos histogramas das classes, priorizando a menor sobreposição das áreas das classes, o que ocasionaria uma confusão na classificação.

O método da distância de Bhattacharya foi executado, em todos os comprimentos de ondas da imagem Hyperion, para a verificação dos melhores conjuntos de bandas, para separação das classes utilizadas. Esta distância estatística utiliza os dados amostrais para a verificação da banda ou conjunto de bandas com a maior distância mínima entre as curvas dos dados amostra. O conjunto de bandas selecionadas foi utilizado nos classificadores Bayesianos Máxima Verossimilhança, Linear de Fisher e Distância Euclidiana Mínima.

Os processos de classificação foram implementados sem *Threshold* e com *Threshold* de 2%, no programa MultiSpec. A análise dos classificadores com *Threshold* objetiva excluir da classificação os pixels com probabilidade de classificação correta menor do que 2%, o que resulta em uma imagem classificada com maior acurácia.

Para comparação dos resultados obtidos nas classificações, foram utilizadas amostras extraídas sobre imagem Ikonos, de alta resolução espacial, e elaborados a partir de interpretação visual juntamente com o trabalho de reconhecimento de campo, e vetorizados no software ARCGIS.

Esses vetores foram utilizados para a construção das matrizes de confusão para a verificação das classes onde existe maior ou menor mistura de informações através do método de Re-substituição e o Leave-One-Out, no Multispec.

Resultados

Com relação à distância entre as classes, os resultados mostraram que o canal que apresentou maior separabilidade foi o de comprimento de onda de 678nm. Neste canal a menor distância apresentada de separabilidade foi entre as classes de Mata e Banana, (5,01), ou seja, este valor representa a separação (diferença) das médias das amostras de treinamento destas duas classes. A maior distância de separabilidade entre as classes

ocorreu entre a classe Sombra e a classe Campo (41,6). A media das distâncias entre as classes neste canal foi de 16.20.

A partir da análise do histograma das amostras de treinamentos das 4 classes, Figura 2, utilizadas para subsidiar as classificações supervisionadas, foi possível verificar a separabilidade entre as amostras coletadas no canal selecionado pela distância estatística de Bhattacharya, comprimento de onda 689 nm.

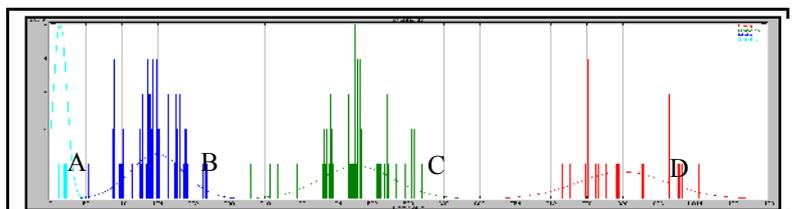


FIGURA 02: Histograma de distribuição dos pixels, no comprimento de onda de 689 nm, das classes Sombra (A), Mata (B), Banana (C), e Campo (D), na imagem do satélite EO-1, sensor Hyperion, data 02/05/2004.

A utilização da distância estatística permitiu um menor tempo de processamento dos dados na classificação, já que foi utilizado um número menor de bandas neste processo, visando otimizá-lo. O resultado da classificação não foi prejudicado, porque foi utilizado o melhor conjunto de 5 bandas possíveis, ou seja, o que permitiu maior separabilidade dos dados.

Analisando-se a separabilidade das classes através de um conjunto de bandas, onde existe uma combinação entre eles, a separabilidade das classes aumentou.

A combinação, contendo 5 bandas, que apresentou melhor separabilidade das classes, utilizou na região do visível os comprimentos de ondas 547 nm (verde) e 628 nm (vermelho); na região do infravermelho próximo o comprimento de onda 730 nm; e na região do infravermelho médio do espectro eletromagnético os comprimentos de ondas 1316 nm e 1720 nm.

Nesta combinação a menor distância mínima foi de 9,83, novamente entre as classes Campo e Banana, e a maior separabilidade foi de 101, entre as classes Campo e Sombra. A média das distâncias nesta combinação foi de 34,02.

No resultado do processo de classificação pelo método da Máxima Verossimilhança os valores de Índice Kappa foram de 100% e 99,7%. No método Leave-One-Out, ocorreu um erro na classificação de um pixel, que era para ser classificado como Sombra e foi classificando como Banana.

A média de probabilidade do método Máxima Verossimilhança foi de 12,7%, conforme a Tabela 2. A classe que mais ocorreu foi a Classe Mata, com 70,33%. Este alto percentual ocorreu devido à área de Mata Atlântica preservada na porção norte da imagem e nos morros da região, onde ocorrem as maiores declividades e altitudes, o que dificulta a intervenção antrópica. O percentual da classe Banana foi de 21,04% predominantemente na parte sul, sudeste e noroeste da imagem, distribuída nas encostas dos vales dos rios da região.

TABELA 2: Distribuição das áreas das classes, método da Máxima Verossimilhança sem *Threshold*.

Classes	No de pixels	Porcentagem	Área (ha)
Campo	2418	4,02	259,99
Banana	12663	21,04	1361,56
Mata	42329	70,33	4551,34
Sombra	2772	4,61	298,05
Total	60182	100,0	6470,95
Probabilidade Média 12,7%.			

No resultado do processo de classificação pelo método Linear de Fisher o valor do Índice Kappa também foi de 100%, não sendo classificado nenhum pixel em uma classe errada.

Assim como o método da Máxima Verossimilhança, o método de classificação Linear de Fisher apresentou como

resultado da classe com maior porcentagem de área, a classe Mata, com um total de 62,72% da área da imagem, sendo 7,61 pontos percentuais menores do que o método anterior. A média de probabilidade do método do classificador de Fisher foi de 41,9%, sendo 29,2 pontos percentuais mais altas do que o método da Máxima Verossimilhança.

TABELA 3: Distribuição das áreas das classes, método de Fisher, sem *Threshold*.

Classes	No de pixels	Porcentagem	Área (ha)
Campo	1762	2,93	189,46
Banana	12724	21,14	1368,12
Mata	37745	62,72	4058,46
Sombra	7951	13,21	854,92
Total	60182	100,0	6470,95

Probabilidade Média 41,9%.

A classificação pelo método da Distância Euclidiana Mínima teve os valores mais baixos de Índices Kappa, 97,2% e 97,0%, para as matrizes de confusão construídas a partir do método de Re-substituição e de Leave-One-Out. No processo de Re-substituição, 11 pixels da classe Banana foram classificados como Campo, 1 pixel da classe Mata foi classificado como Banana e 1 pixel da classe Campo também foi erroneamente classificado como Banana. No processo Leave-One-Out os valores ficaram praticamente os mesmos, exceto por 1 pixel a mais da classe Campo que foi classificado como Banana.

Algumas áreas classificadas como Campo pelo método de Fisher foram classificadas como Banana pelo método da Distância Euclidiana Mínima e, outras áreas que foram classificadas como Banana, foram classificadas como Mata. Em relação a estas diferenças apresentadas verificou-se que as áreas classificadas pelo método de Fisher estão corretas. Provavelmente porque método

Distância Euclidiana Mínima não utiliza a variância no processo de classificação (MÜLLER, DALMOLIN, ARAKI, 1999).

TABELA 4: Distribuição das áreas das classes, método da Distância Euclidiana Mínima.

Classes	No de pixels	Porcentagem	Área (ha)
Campo	3042	5,05	327,09
Banana	10847	18,02	1166,30
Mata	37777	62,77	4061,90
Sombra	8516	14,15	915,67
Total	60182	100,0	6470,95

Os resultados do Índice Kappa no processo de classificação pelo método da Máxima Verossimilhança com *Threshold* de 2% foram de 94,6% e 93,5%. A imagem resultante tem uma menor área classificada neste processo, porque são eliminadas as áreas com probabilidade de acerto menor de 2 %, o que resulta numa imagem com maior precisão para o usuário.

A classe Mata foi a área mais representada na imagem, com 40,57 %, a classe Banana ficou com apenas 6,15 %. A classe Campo foi a menos representada com 0,91% e a classe Sombra representa 3,40 % da imagem. No total 48,98 % da imagem não foi classificada por causa *Threshold* de 2% (Tabela 5).

TABELA 5: Distribuição das áreas das classes, método da Máxima Verossimilhança com *Threshold* de 2%.

Classes	No de pixels	Porcentagem	Área (ha)
Campo	546	0,91	58,71
Banana	3699	6,15	397,73
Mata	24417	40,57	2625,39
Sombra	2044	3,40	219,78
<i>Não classificada</i>	29476	48,98	3169,35
Total	60182	100,00	6470,95

No resultado do processo de classificação pelo método de Fisher com *Threshold* de 2%, o valor de Índice Kappa foi de 94,0%. Nenhum pixel foi substituído em classe errônea, entretanto 22 pixels das amostras não foram classificados, sendo 4 na classe Campo, 11 na classe Banana, 6 na classe Mata e 1 na classe Sombra. As classes Mata e Banana tiveram um aumento na área classificada em relação ao método anterior e, na classe Campo, ocorreu uma diminuição, ficando cada classe com 48,59%, 7,80 % e 0,72 % respectivamente.

Com relação ao método Máxima Verossimilhança ocorreu um aumento da área classificada da imagem. O classificador de Fisher, utilizando o mesmo limiar de 2 % para exclusão dos pixels, deixou de classificar 31,88%, desta forma o método de Fisher classificou 17,10% a mais da imagem Hyperion do que o método Máxima Verossimilhança. Este aumento ocorreu porque o classificador de Linear de Fisher é uma simplificação do método Máxima Verossimilhança, fazendo com que o mesmo seja menos restritivo no momento da classificação.

O valor do Índice Kappa, para o método de classificação da Distância Euclidiana Mínima, comparado com os dados extraídos da imagem de alta resolução espacial, foi de 0,2238 e a Exatidão Global foi de 0,5952. Estes valores foram os mais baixos dentre os classificadores avaliados. O total de área que foi classificada corretamente foi de 1208,22ha, dos 2033,34ha classificados.

O vetor da classe Banana representava 325,56ha, que é equivalente a aproximadamente 3617 pixels da imagem Hyperion. Desta área total, 150,84ha foram classificados corretamente, o que corresponde a aproximadamente 1676 pixels. A classe Banana apresentou grande confusão na classificação com a classe Mata, já que mais da metade das áreas vetorizadas como Banana foram classificados como Mata, 162,81ha, o que corresponde a aproximadamente 1808 pixels. Tisot *et al*,2007 destaca a importância da imagem Hiperespectral para a discriminação de alvos agrícolas que apresentam diferenças espectrais sutis, mas devido ao classificador que utiliza o método da Distância

Euclidiana Mínima ser muito limitado, o mesmo não consegue detectar tais diferenças e acaba rotulando alguns pixels erroneamente.

O valor do Índice Kappa para o método de classificação Linear de Fisher foi de 0,2612 e a Exatidão Global da classificação foi de 0,5970. Estes valores são maiores do que os valores da classificação pelo método da Distância Euclidiana Mínima. O total da área classificada corretamente por este método foi de 1213,92 ha, dos 2033,34 ha classificados.

A classe Banana foi a única classe que teve um aumento de área classificada corretamente, esse aumento foi de 33,77ha. As classes Campo, Mata e Sombra tiveram uma diminuição da sua área classificada corretamente, essa diminuição foi de 9,34 ha, 16,78 ha e 1,95 ha respectivamente.

O método de classificação de Fisher aplicado com um *Threshold* de 2% apresentou uma significativa melhora nos seus Índices de acurácia. Sendo que o valor de Kappa ficou 0,3456 e o valor da Exatidão Global foi de 0,6951. O total de área que foi classificada corretamente foi de 858,28 ha, dos 1234,73 ha classificados.

A classe Banana teve 119,01 ha classificados corretamente, dos 192,71 ha classificados na imagem. A Classe Mata teve 703,87 ha classificados corretamente, dos 957,89 ha. Os erros de omissão e de comissão mais baixos foram os das classes Sombra e Mata.

Analisando o erro de comissão mais baixo, que foi o da classe Mata, 0,11, significa que 11% da área desta classe deveria estar distribuído em outras classes e não na classe Mata. E analisando o erro de omissão da mesma classe, 0,27, significa que 27% de áreas que estão distribuídas nas outras classes deveriam ter sido classificadas como Mata.

O método de classificação de Máxima Verossimilhança apresentou o mais alto resultado de Exatidão Global e Índice Kappa, entre os métodos sem *Threshold*. O valor de Kappa foi de 0,3218 e o valor da Exatidão Global foi de 0,6911.

A classe que teve o menor erro de omissão e de comissão foi a classe Mata, tendo 0,25 de erro de omissão e 0,14 de comissão. A classe Banana teve o segundo menor erro de omissão, 0,42 e o maior erro de comissão, 0,70.

O método de classificação de Máxima Verossimilhança com *Threshold* de 2% apresentou o melhor resultado de Exatidão Global e também de Índice Kappa, entre todos os métodos utilizados. O valor de Kappa foi de 0,4885 e o valor da Exatidão Global foi de 0,7755. Entre todos os métodos de classificação analisados, este foi o que obteve os maiores valores de Índice Kappa e de Exatidão Global.

A classe que teve o menor erro de omissão e de comissão foi a classe Mata, tendo 0,17 de erro omissão e 0,07 de comissão. A classe Banana teve o terceiro menor erro de omissão, 0,26 e também o terceiro menor erro de comissão, 0,55. A classe Sombra teve o maior erro de comissão, 0,71 e o segundo menor erro de omissão, 0,20.

Analisando-se os resultados de todos os processos de classificação, verifica-se uma melhora na acurácia quando é utilizado o limiar de exclusão (*Threshold* 2%). O índice Kappa aumentou 0,08, para o classificador de Fisher e 0,16 para o de Máxima Verossimilhança. A Exatidão Global também aumentou, sendo o aumento de 0,09 para Fisher e 0,11 para a Máxima Verossimilhança.

O processo de classificação Máxima Verossimilhança com o *Threshold* de 2% foi o método que classificou a menor quantidade de área das amostras vetorizadas, 864,22 ha, já o método Linear de Fisher com *Threshold* de 2% classificou 370,51 ha a mais, totalizando 1234,73 ha.

A grande diferença entre os valores de acurácia, principalmente nos Índices Kappa apresentados para a análise com os dados das próprias amostras de treinamento do classificador comparado com os dados extraídos da imagem de alta resolução espacial, deve-se principalmente pela diferença de resolução espacial das imagens, Hyperion (30 m) e IKONOS (1 m).

A maior precisão dos dados vetoriais permitiu identificar outras classes de uso do solo, que não foram contabilizadas no processo de classificação, mas, que foram identificadas no processo de vetorização, como solo exposto e outros cultivos. Como as amostras de treinamento para os classificadores foram extraídas da imagem Hyperion, algumas amostras podem ter sido contaminadas por outras classes de uso do solo não contabilizadas no processo de classificação.

Como os vetores destas outras áreas não foram utilizados na avaliação, presume-se que grande parte das áreas que foram classificadas erroneamente, pertençam a estas classes. Como o Índice Kappa é de avaliação mais criteriosa, e leva em consideração não somente os acertos, mas também os erros e a própria distribuição dos erros entre as classes, os baixos valores apresentados pelas classificações demonstram que existe um elevado grau de erro, e que este erro está distribuído em todas as classes.

TABELA 6: Índices de acurácia para os métodos de classificação, utilizando vetoriais extraídos da imagem Ikonos.

Classificadores	Índices de acurácia	
	Kappa	Exatidão global
Máxima verossimilhança	0,3218	0,6611
Máxima verossimilhança c/ <i>Threshold</i> 2%	0,4885	0,7755
Distância euclidiana mínima	0,2238	0,5942
Fisher	0,2612	0,5970
Fisher c/ <i>Threshold</i> 2%	0,3456	0,6951

Conforme classificação sugerida por Vieira & Soares (2008), os valores admitidos de índice Kappa entre 0,2 e 0,4 são razoáveis e, os valores entre 0,4 e 0,6 são bons, em uma classificação que varia entre péssimo a excelente. Desta maneira somente o método da Máxima Verossimilhança com *Threshold* de

2% apresentou um bom resultado de Índice Kappa, todos os outros métodos utilizados são classificados como razoáveis conforme esta classificação.

Oliveira (2009) realizou um estudo similar utilizando imagem HRG-SPOT 5, com os classificadores, Máxima Verossimilhança Gaussiana, Distância Euclidiana Mínima, Battacharya e Iseseg, para o município de Três Cachoeiras. Neste trabalho foram utilizadas as seguintes classes: Floresta Ombrófila Densa, Formação Florestal Secundária, Florestamento, Bananais, Solo Exposto, Campo e Sombra.

Os resultados obtidos por Oliveira (2009) comparando com um fragmento vetorizado da mesma imagem utilizada nos processos de classificação são apresentados na Tabela 7. O resultado do classificador Máxima Verossimilhança Gaussiana, foi de 0,5054 para o Índice Kappa e de 61,78 % para a Exatidão Global. Comparando este resultado com o encontrado na classificação Máxima Verossimilhança com *Threshold* de 2% para as 5 bandas selecionadas da imagem Hyperion utilizada neste trabalho, houve uma melhora na Exatidão Global de 15,77 pontos percentuais. Mesmo comparando com o melhor resultado de Oliveira (2009), que foi o do Bhattacharya, o valor da Exatidão Global encontrado na classificação Máxima Verossimilhança com *Threshold* de 2% foi 8,42 pontos percentuais superior.

Os melhores valores de Exatidão Global, se comparados aos encontrados por Oliveira (2009), são resultado das vantagens da utilização de dados hiperespectrais. Dados hiperespectrais permitem uma melhor discriminação dos alvos, visto que tem a capacidade de compor uma imagem com resolução espectral com um nível de detalhamento bem próximo dos verificados em espectrorradiômetros de campo ou laboratório.

Como o procedimento de extração de feições pelo método de Bhattacharya selecionou as melhores bandas para separação destas classes, com base nas características das amostras de treinamento extraídas da imagem, para serem utilizadas nos processos de classificação, otimizou-se o resultado da classificação. A

BISSO, A.B.A. et al. Análise dos Classificadores Bayesianos utilizando ...

eliminação das bandas, cujos comprimentos de ondas diferenciam menos as classes, auxilia no melhor resultado da exatidão da classificação, já que essas bandas fariam com que os resultados das funções de decisão dos classificadores fossem mais baixos, possibilitando uma maior confusão no processo de classificação.

TABELA 7: Tabela dos Índices Kappa e Exatidão Global dos classificadores com dados HRG – SPOT 5, para o município de Três Cachoeiras. Fonte: Oliveira (2009).

Classificadores	Índices de acurácia	
	Kappa	Exatidão Global
Máxima Verossimilhança Gaussiana	0,5054	61,78%
Distancia Euclidiana Mínima	0,3647	50,11%
Bhattacharya	0,5753	69,13%
Isoseg	0,5194	64,33%

Como já discutido anteriormente, devido à diferença de resolução espacial e espectral das imagens analisadas e, principalmente, pela existência de outras classes de uso do solo, identificadas na imagem de alta resolução, o Índice de Exatidão Global representa melhor a qualidade dos resultados das classificações, já que o mesmo verifica somente os valores das áreas corretamente classificadas.

Desta maneira podemos dizer que o dado hiperespectral contribui para a melhoria da classificação, e que os baixos valores de Índice Kappa não representam de forma adequada este avanço. Provavelmente estes baixos valores do Índice Kappa são ocasionados por pequenas áreas de outras classes de uso do solo, que se encontram inseridas em regiões com um predomínio diferente de uso e ocupação do solo. Devido a baixa resolução espacial do sensor Hyperion, não se pode evidenciar esta diferença. Fazendo com que as pequenas áreas fossem erroneamente rotuladas no processo de classificação.

Para melhoria em uma análise futura, sugere-se utilizar os próprios dados extraídos da imagem de alta resolução espacial, para efetuar o treinamento e avaliação dos classificadores, utilizando-se de 50% dos dados vetoriais para treinamento e os outros 50% para validação da classificação. Desta maneira diminuiríamos a probabilidade da ocorrência de amostras “poluídas” com outras classes de uso do solo, o que possivelmente aumentaria a acurácia dos classificadores, já que trabalharíamos com dados numa mesma escala.

Considerações finais

Com base nos resultados obtidos neste estudo, podemos destacar as seguintes conclusões em relação a análise das áreas de Mata Atlântica e de cultivo de bananas no Litoral Norte gaúcho do estado do Rio Grande do Sul.

Em relação aos testes dos classificadores Bayesianos, verificou-se que o método Máxima Verossimilhança com *Threshold* de 2% apresentou o melhor resultado para a separabilidade entre as classes de Banana e Mata Atlântica. Apresentou os maiores valores de Índice Kappa, de Exatidão Global e também um dos menores erros de omissão e comissão, demonstrando o seu alto potencial de classificação. A utilização do *Threshold* no processo de classificação representou um ganho na acurácia.

Os dados hiperespectrais demonstraram ser uma boa alternativa para análise ambiental, trazendo outra forma de abordagem e análise. Mesmo utilizando-se de métodos tradicionais de classificação, os resultados atingiram um bom desempenho com a utilização do critério de seleção de feições, Bhattacharya.

Recomenda-se avaliar o poder de discriminação do classificador SAM, *Spectral Angle Mapper*, como forma de validar o potencial de separação de classes de uso e ocupação do solo utilizando dados hiperespectrais, e compará-los com os resultados encontrados nos classificadores tradicionais.

Referências bibliográficas

FELDE, G.W; ANDERSON, G.P; COOLEY, T.W; MATTHEW, M.W.; ADLERGOLDEN, S.M.; BERK, A.; LEE, J. **Analysis of Hyperion data with the FLAASH atmospheric correction algorithm.** In: INTERNATIONAL GEOSCIENCE AND REMOTE SENSING SYMPOSIUM, n.3, 21-25 jul. 2003, [S.I.]. Proceedings... v.1, p.90-92.

GALVÃO, L.S; FORMAGGIO, A. R; BREUNIG, F. M. **Relações entre índices de vegetação e produtividade de soja com dados de visada fora do nadir do sensor Hyperion/EO-1** In: XIV Simpósio Brasileiro de Sensoriamento Remoto, Natal, Anais, INPE, 25-30 abril 2009, p.1095-1102.

MÜLLER, S.I.M.G., DALMOLIN, Q., ARAKI, H. **Comparação entre os métodos de Máxima Verossimilhança, Distância Mínima e o método de Fisher para reconhecimento de padrões em imagens coloridas** Bol. Ciênc. Geod., Curitiba, v. 4, p.3-13, 1999.

NOGUCHI, M.T.G. **Um Sistema Baseado em Regras Fuzzy para Classificação Supervisionada de Imagens Multiespectrais de Alta Resolução.** 2004. 150 f. Dissertação de Mestrado. Programa de Pós-Graduação em Informática, Setor de Ciências Exatas, Universidade Federal do Paraná, Curitiba – PR, 2004.

OLIVEIRA, G. **Integração de dados HRG-SPOT5 e SRTM para espacialização de bananais em área do bioma mata Atlântica no município de Três Cachoeiras, RS.** 2009. 83 f. Monografia (graduação), Instituto de Geociências, Curso de Geografia bacharelado, Universidade Federal do Rio Grande do Sul, Porto Alegre - RS, 2009.

BISSO, A.B.A. et al. Análise dos Classificadores Bayesianos utilizando ...

SCHRAMM, V.F; VIBRANS, A.C. **Uso de imagens hiperespectrais (EO-1 Hyperion) para detalhamento da detecção das formações florestais na bacia do Itajaí** In: Dynamis revista tecno-científica , n.1, out-dez/2007,v.13, p.59-69.

SENSORIAMENTO Remoto Hiperespectral. INFOGEO. Curitiba, PR: Espaço GEO, v.6, n.33, maio/jun. 2004.

TISOT, D.A.A; FORMAGGIO, A.R; RENNÓ, C.D; GALVÃO, L.S. **Eficácia de dados Hyperion/eo-1 para identificação de alvos agrícolas: comparação com dados ETM+/Landsat-7** In: Eng. Agríc., Jaboticabal, n2 maio/ago, 2007, v.27, p.511-519.

VIEIRA, C.A.O; SOARES, V.P. **Projeto UFV Geocapacitar: Processamento Digital de Imagens.** 2008. Universidade Federal de Viçosa – MG, 2008.

Recebido em março de 2011
Aceito em agosto de 2011